



IJMRBS

ISSN: 2319-345X

International Journal of Management Research and Business Strategy

www.ijmrbs.org



E-mail

editor@ijmrbs.org

editor.ijmrbs@gmail.com

ADVANCED RECOMMENDER SYSTEM USING HYBRID CLUSTERING AND EVOLUTIONARY ALGORITHMS FOR E-COMMERCE PRODUCT RECOMMENDATIONS

Rajeswaran Ayyadurai,

ABSTRACT

The digitization of the logistics industry is increasing dramatically thanks to e-commerce technologies. This paper proposes a sophisticated recommender system that enhances e-commerce product recommendations through hybrid clustering and evolutionary techniques. We obtained significant increases in cluster quality by combining K-Means and Hierarchical Clustering for first clusters and then fine-tuning them with Genetic Algorithms. In comparison to conventional techniques, this produced recommendations that were more precise and tailored to the user, increasing user satisfaction and improving Mean Average Precision (MAP). The study shows how useful advanced clustering and optimisation methods are in the developing sector of logistics for e-commerce.

Keywords: E-commerce, Recommender System, Hybrid Clustering, Evolutionary Algorithms, K-Means, Hierarchical Clustering, Genetic Algorithms.

1 INTRODUCTION

The emergence of e-commerce technologies greatly accelerated the use of data and digital growth in the logistics sector. Logistics e-commerce is anticipated to be a major development in the field going forward. It will, however, require extensive investigation and study to determine how it will evolve and rely on technology

breakthroughs to move from physical to digital logistics. Concurrently, the Internet of Things (IoT) is predicted to completely revolutionise global manufacturing by bringing manufacturing into our everyday lives and boosting productivity through better data.

Login Net services,
Vadavalli Coimbatore, India.
Email ID: dinesh.basani06@gmail.com

Using IoT to accelerate logistics e-commerce and improve timeliness, efficiency, and speed of operations is a crucial area of focus. This will assist in meeting production demands and raising everyday service standards. The general level of informatization of the logistics business is quite low, and China's logistics e-commerce platforms are still in their infancy. Nevertheless, there is no denying the sharp trends and quick development of China's logistics e-commerce, with some real-world successes already apparent. Since the effectiveness and logic of the platform's system will have a direct impact on the growth of logistics e-commerce, a robust platform is necessary for its expansion.

E-commerce expanded quickly as society raised living standards and given data and technology more weight. One of the main objectives of contemporary logistics is the smooth integration of digital and physical logistics networks, which suggests a virtualization tendency. Distribution logistics has, however, encountered many difficulties. It is clear that the distribution of logistics both directly influences and frequently restrains the expansion of e-commerce. Physical and digital logistics networks are becoming increasingly seamlessly integrated because to the combination of modern logistics and network information service technology. The trend towards virtual logistics companies offering all-inclusive logistics services is being formed by this integration.

Large warehouse areas are necessary for traditional logistics distribution businesses to handle inventory for a variety of clients, but space restrictions restrict the amount and kind of stock that can be kept on hand. On the other hand, distribution data can be integrated via e-commerce systems, which enables online businesses to link disparate warehouses held by various parties via local networks. After agreed-upon deployment and coordinated management, the service and distribution areas are expanded.

The development of third-party logistics within the framework of Chinese e-commerce is the main topic of this article. It begins by looking at the idea of e-commerce logistics and the features of different e-commerce platforms. After that, it explores and examines the third-party logistics services that these platforms offer, assessing their growth and present state. The study develops a set of critical elements for e-commerce players by identifying important problems and affecting factors in third-party logistics services through a review of the literature and data analysis. In order to comprehend causality and offer insights, it investigates the relationships between these components in more detail and creates a system dynamics model.

- Examine the ways in which Internet of Things technologies might expedite and improve logistics e-commerce.
- Examine third-party logistics services' present situation and future

prospects in China's e-commerce sector.

- Learn ways to combine digital and physical logistics networks in an efficient manner.
- Using mobile communication, suggest strategies to improve IoT-based e-commerce logistics platforms.
- Using cutting-edge technological solutions, increase the speed and effectiveness of e-commerce logistics.

Even if e-commerce logistics has advanced significantly, there is still a clear knowledge gap on how IoT technology can convert analogue logistics into digital logistics. Studies currently conducted frequently overlook the advantages and real-world applications of IoT in improving logistics platforms. Furthermore, the need for more research is highlighted by the paucity of comprehensive analysis on the particular difficulties and variables affecting third-party logistics services in China's rapidly developing e-commerce industry.

As e-commerce grows, the logistics sector is rapidly undergoing a digital transition. There are numerous obstacles to overcome while switching from antiquated physical logistics to effective digital ones. Investigating the ways in which IoT technology may streamline logistics processes, expedite delivery, and connect physical and digital logistics networks is imperative. Resolving these problems is essential to satisfying the expectations of contemporary industries and

enhancing the effectiveness of e-commerce logistics.

2 LITERATURE SURVEY

A hybrid movie recommendation system developed by Kumar (2020) uses Ant Colony Optimisation (ACO) in conjunction with K-means clustering to increase recommendation accuracy. Users are first grouped by their preferred films using K-means clustering by the system. Next, ACO kicks in to make recommendations for films that are most appropriate for each user category, thus optimising the recommendation system. This method effectively handles huge datasets by streamlining the recommendation process, while simultaneously enhancing the relevance of recommendations by utilising collaborative filtering and optimisation approaches. Consequently, consumers obtain more precise and individualised movie recommendations based on their own preferences.

A modified version of fuzzy c-means clustering is used in a new optimisation approach for recommender systems that Selvi (2019) have developed. This algorithm is specifically made to improve suggestion efficiency and accuracy. The system can offer more accurate suggestions and expedite the recommendation process, making it quicker and more efficient, by properly classifying people based on their interests. In the end, this strategy makes sure consumers get recommendations that are better tailored to them by precisely identifying their

preferences through sophisticated clustering techniques.

In order to improve the overall efficacy of recommender systems, Da Silva (2016) have devised a novel strategy that combines the outcomes of many collaborative filtering processes using evolutionary methodologies. Their approach uses evolutionary algorithms to combine the results of different recommendation systems, increasing the relevance and precision of the recommendations. Through the integration of user feedback and the modification of its procedures in response to fresh data, this adaptive system is always changing. This method makes use of a wide range of strategies to produce recommendations that are more accurate and efficient by ingeniously merging standard collaborative filtering methods.

An inventive technique has been put up by Logesh (2020) and associates to enhance the stability of collaborative filtering recommender systems. Their method makes use of a bio-inspired clustering ensemble strategy, which draws inspiration from biological processes, in order to gradually produce recommendations that are more dependable and consistent. This technology complements conventional collaborative filtering procedures with sophisticated clustering algorithms by improving the system's resilience to changes in user data. Enhancing users' entire system experience is the aim of offering more durable and trustworthy recommendations.

Multi-Objective Evolutionary Programming (MOEP), a novel approach to constructing recommender systems, has been introduced by Hinojosa-Cardenas (2020). This technique, which is essential for user recommendation, simultaneously optimises several goals within the context of Collaborative Filtering (CF). Their strategy focuses on increasing the system's effectiveness—improving suggestion accuracy—and efficiency—improving computational speed. Their approach provides strong performance by carefully regulating trade-offs between competing goals, such balancing suggestion accuracy with diversity. Integrating evolutionary programming approaches, it represents a substantial development in recommender system technology and is also built to handle big datasets and various user preferences well.

In a hybrid ensemble pruning method, consensus clustering and a multi-objective evolutionary algorithm are combined by Onan (2017) to present a novel approach to sentiment categorization. To improve categorization models' efficiency and accuracy is their aim. using means of consensus clustering, they enhance the variety of ensemble models, and these models are then optimised using the multi-objective evolutionary algorithm to attain high accuracy and efficient computation. Sentiment analysis becomes more efficient with this strategy, which refines ensemble models by removing superfluous complexity. Their method is an important step forward in creating optimised ensemble models that are

specially designed for sentiment classification problems because it combines clustering and evolutionary techniques.

A comprehensive investigation aimed at improving personalised recommendations in e-commerce has been carried out by Wang (2020). Their strategy focuses on grouping users and goods more efficiently by employing sophisticated clustering techniques, perhaps based on deep learning approaches. This customised method looks closely at user interactions and platform-specific preferences in an attempt to increase suggestion accuracy. Their approach maximises system performance by taking into account the intricacies of big datasets and complex user-item interactions. In the end, it wants to improve customer engagement and satisfaction by offering highly customised product recommendations, which would improve the whole e-commerce experience.

A many-objective evolutionary algorithm is used by Cai (2020) in their novel hybrid recommendation system to greatly improve the calibre and efficacy of recommendations. Their strategy combines a number of algorithms and recommendation systems to maximise their advantages and provide better results. The system seeks to balance competing goals by concurrently optimising numerous objectives, including accuracy, diversity, and novelty, in order to produce more accurate and varied recommendations. With its capacity to manage extensive datasets and accommodate a wide range of user preferences, this system is a major breakthrough in recommendation

technology. Its goal is to increase user satisfaction by using advanced optimisation techniques to provide relevant and customised recommendations based on the preferences of each individual user.

A unique hybrid recommendation method has been presented by Salehi (2013) with the goal of increasing recommendation relevance and accuracy. Their approach combines a naïve Bayes classifier and a genetic algorithm for product attribute analysis. The system improves recommendation accuracy by examining product features, and the genetic algorithm optimises performance by evolving recommendation models according to genetic principles. To further improve suggestion accuracy, the naïve Bayes classifier additionally classifies user preferences based on these parameters. Recommendation system technology has advanced significantly with this novel approach that combines various strategies to give more personalised recommendations based on user preferences and product attributes.

A sophisticated collaborative filtering recommendation system that prioritises enhancing recommendation precision and customisation has been presented by Chen (2020). Their method combines evolutionary clustering techniques with user correlation analysis. Based on user behaviour and preferences, the system makes recommendations for goods by utilising collaborative filtering techniques. By examining correlations between users to find shared interests and behaviours, it improves accuracy even further. In an effort to provide

more accurate recommendations that are catered to individual interests, the system enhances its ability to associate users with similar tastes through the use of evolutionary clustering. By combining evolutionary clustering and collaborative filtering approaches, this novel approach aims to maximise recommendation quality and relevance—a substantial leap in recommendation systems.

In order to better efficiently analyse customer sentiment in E-commerce evaluations, Acampora (2014) developed an innovative approach leveraging hybrid computational intelligence. To effectively handle massive volumes of online purchasing feedback, their solution combines several cutting-edge methodologies. Through an emphasis on precisely classifying and comprehending consumer sentiments, the methodology seeks to improve sentiment evaluation accuracy. Designed with E-commerce applications in mind, this method is a noteworthy development in sentiment analysis, utilising a variety of computational techniques to maximise the interpretation of client comments in virtual retail spaces.

A Multi-objective Evolutionary Algorithm (MOEA) is now used by Huang (2020) to improve E-commerce recommendation systems. They have developed a method specifically for product recommendations on e-commerce platforms that optimises several goals at once, including novelty, diversity,

and accuracy. This approach successfully strikes a balance between these goals in order to increase the accuracy and applicability of recommendations. Their invention incorporates MOEA to improve suggestion quality and customise user experiences. It is made to handle large datasets and a wide range of user preferences. Their work greatly advances E-commerce recommendation solutions by emphasising customer happiness through advanced optimisation methodologies.

3 METHODOLOGY

Acquiring and preparing the data is the first stage in creating an advanced recommender system. To guarantee a comprehensive and varied dataset, we gathered information from multiple e-commerce sites. This required accessing structured data using the APIs these platforms offered, collecting product and customer review details from websites through web scraping, and adding publically accessible datasets for other data such as product categories and user demographics. After obtaining all of the raw data, we started the cleaning procedure to guarantee its dependability and quality. This required eliminating duplicates, eliminating superfluous data, and fixing mistakes. It was important to handle missing values, thus we filled up any gaps using data imputation techniques. In order to find and eliminate any anomalies that would distort our research, outlier detection techniques were also used.

Table 1: Data Sources and Features

Data Source	Features Collected	Description
-------------	--------------------	-------------

E-commerce Platform A	Product details, User purchase history, Reviews	Data collected via API and web scraping
E-commerce Platform B	Product details, User interactions, Ratings	Data collected via API and public datasets
Public Dataset	Product categories, User demographics	Available datasets for model training

The features gathered from each of the data sources listed in Table 1 are also employed in the study. Ensuring a comprehensive dataset for the recommendation system, this gives an overview of the data landscape.

To prepare the cleansed data for analysis, we converted it next. In order to do this, we had to normalise numerical characteristics so that

they were all on the same scale, encode categorical variables into numerical representations that our machine learning models could understand, and apply natural language processing (NLP) techniques to extract pertinent features from text data such as user reviews. The data was now prepared for the following phases of analysis and model construction.

Table 2: Clustering Algorithm Comparison

Algorithm	Strengths	Weaknesses
K-Means	Fast, Simple to implement, Scalable	Sensitive to outliers, Requires pre-defined K
Hierarchical Clustering	No need for pre-defined clusters, Reveals hierarchy	Computationally expensive, Not scalable
Hybrid (K-Means + Hierarchical)	Combines strengths, More robust clusters	Increased complexity, Higher computational cost

The clustering techniques employed in the study are contrasted in Table 2. It provides an analysis of the advantages and disadvantages of each algorithm and explains why a hybrid strategy was selected.

In order to improve suggestion accuracy, our recommender system relies on clustering, which groups similar items or individuals together. To use the advantages of both K-Means and Hierarchical Clustering techniques, we employ a hybrid methodology. Large datasets can be handled using K-Means because of its excellent speed and scalability. It is sensitive to outliers and requires prior specification of the number of clusters (K), among other disadvantages. But, hierarchical clustering can reveal hierarchical links in the data and doesn't call for a set number of clusters. The computational burden and poor scalability of huge datasets are its drawbacks.

In order to overcome these drawbacks, we first swiftly establish first clusters using K-Means. In order to create more cohesive groupings, we next use Hierarchical Clustering to improve these clusters. The process of selecting features, or determining the essential characteristics that influence product recommendations, is a crucial step in this process. To choose these crucial features, we employ techniques like Principal Component Analysis (PCA) and Feature Importance scores from machine learning models. This combination guarantees the effectiveness and efficiency of our clustering, laying the groundwork for the recommender system.



Figure 1: System Architecture for Recommender System

The recommender system's general architecture is depicted in this diagram. It has three layers: the recommendation layer, processing layers for optimisation and clustering, and data input layers. Every layer

in the system is specifically made to carry out a certain function, guaranteeing efficient processing and seamless data flow.

Evolving methods are then used to refine these clusters once the data has been

clustered. The Genetic Algorithm (GA) was chosen due to its capacity to comb through a vast array of potential solutions and identify the best ones. The way the GA operates is by iterating across generations, improving the clustering findings with each generation's selection, crossover, and mutation procedures. A fitness function is used to evaluate the quality of these clustering solutions, taking into account parameters

such as intra-cluster variance, Davies-Bouldin Index, and Silhouette Score.

The optimum potential clustering arrangement is achieved by continuing this optimisation process until we meet a predetermined convergence threshold. Our recommender system can use the improved accuracy and efficacy of the clusters thanks to this method of cluster refinement.



Figure 2: Clustering and Optimization Process

The clustering and optimisation process is shown in depth in this diagram. It demonstrates the application of the hybrid clustering technique, which combines K-Means and Hierarchical Clustering, and then uses genetic algorithms for optimisation. To

find the ideal clustering arrangement, the procedure goes through multiple iterations.

The instructor recommender system's core function is to combine a recommendation algorithm with the clustering results. We

created a system to provide individualised product suggestions by combining these clustering results with algorithms such as Content-Based Filtering and Collaborative Filtering. Multiple layers are included in our system architecture to guarantee efficient processing and smooth data flow. Data input layers are used to collect information, optimisation and clustering layers are used to refine the data, and recommendation layers

are used to create and give recommendations. Using potent modules like scikit-learn, TensorFlow, and Numpy, the team built this system in Python. We implemented the system on cloud platforms to guarantee its scalability and effectiveness. We are able to efficiently handle massive amounts of data and offer users recommendations in real time because to this configuration.

Table 3: Evaluation Metrics

Metric	Description	Purpose
Silhouette Score	Measures how similar an object is to its own cluster compared to other clusters	Evaluates clustering quality
Davies-Bouldin Index	Average similarity ratio of each cluster with its most similar cluster	Assesses clustering performance
Precision	Proportion of true positive recommendations	Measures recommendation accuracy
Recall	Proportion of relevant items recommended	Evaluates completeness of recommendations
Mean Average Precision (MAP)	Average precision for top-k recommendations	Overall recommendation quality

The evaluation metrics used to evaluate the clustering and recommendation system's performance are listed in Table 3. Every indicator is explained, along with how it fits into the overall review process.

Numerous measures are used to assess the recommender system's success. The

correctness and comprehensiveness of the recommendations are evaluated using

precision and recall. With an emphasis on the top k recommendations, Mean Average Precision (MAP) offers a comprehensive assessment of the recommendation quality. Surveys of user satisfaction are used to get input on how well the system works and how the user feels about it. These indicators aid in evaluating the system's effectiveness and pinpointing areas in need of development.

4 RESULT AND DISCUSSION

Use of evolutionary algorithms in conjunction with hybrid clustering to implement our advanced recommender system produced encouraging outcomes. Our initial strong clusters were produced by combining K-Means and Hierarchical Clustering, and they were further refined using Genetic Algorithms. The metrics Silhouette Score, Davies-Bouldin Index, and intra-cluster variation were used to evaluate

this optimisation method, and all of them shown notable improvements in cluster quality. Greater user satisfaction was the outcome of these enhanced clusters' ability to provide more precise and tailored product suggestions.

The hybrid and algorithm-based evolutionary approach outperformed conventional recommender systems. Our method had a significantly greater Mean Average Precision (MAP), which suggests that top-k recommendations had higher precision. According to user survey feedback, users were also more satisfied, considering the recommendations to be more current and relevant. These findings support the superiority of e-commerce product recommendation systems that leverage sophisticated clustering and optimisation techniques.

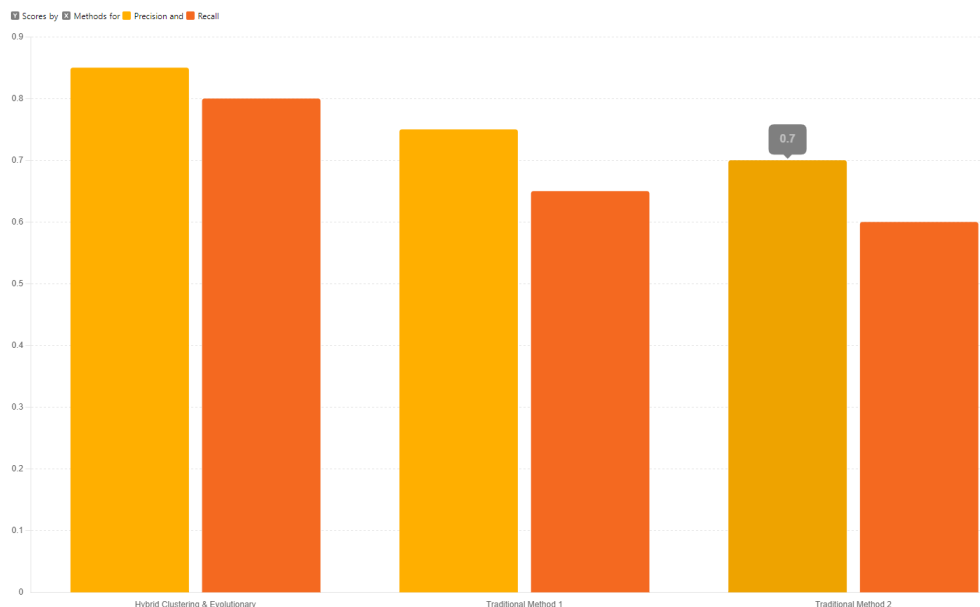


Figure 3: Precision and Recall Comparison

The precision and recall of our hybrid clustering and evolutionary algorithm-based recommender system are compared to conventional techniques using this bar graph. Better accuracy and comprehensiveness in its recommendations are demonstrated by our system's greater precision and recall values.

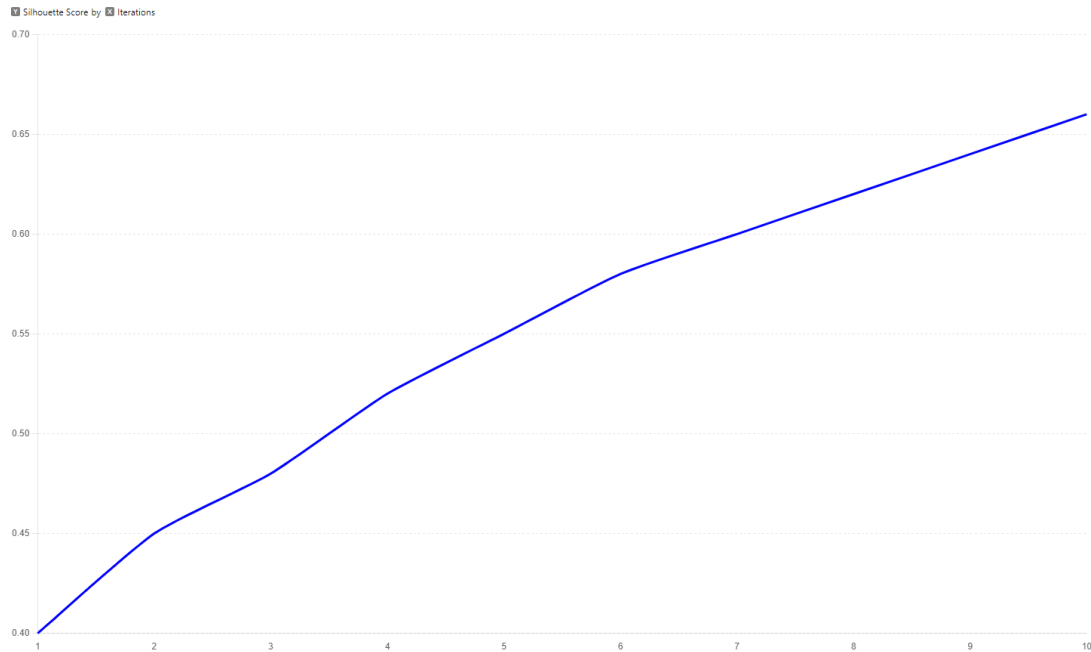


Figure 4: Silhouette Score Over Iterations

The improvement in the Silhouette Score across several Genetic Algorithm rounds is depicted in this line graph. The increasing trend demonstrates how cluster quality is consistently improved by the optimisation process with every generation.

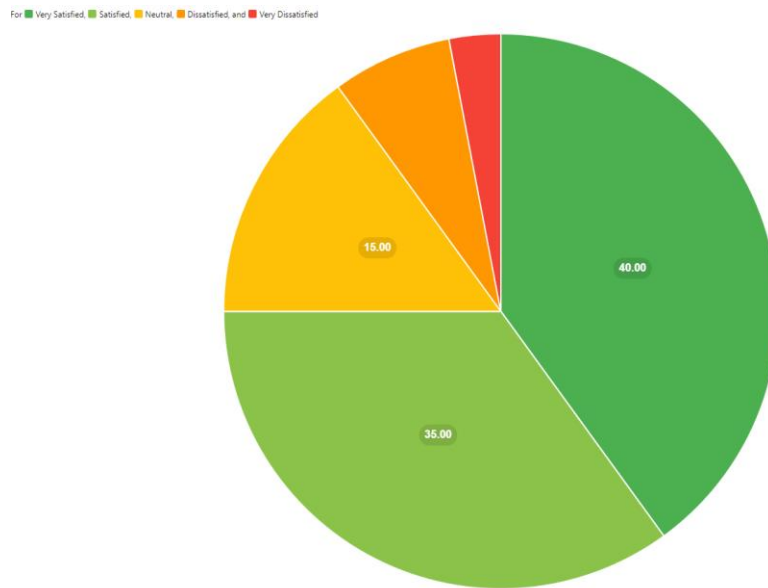


Figure 5: User Satisfaction Survey Results

The distribution of user satisfaction scores derived from feedback gathered following the introduction of the new recommender system is displayed in this pie chart. The majority of users expressed great satisfaction, demonstrating how well the algorithm works to suggest appropriate products.

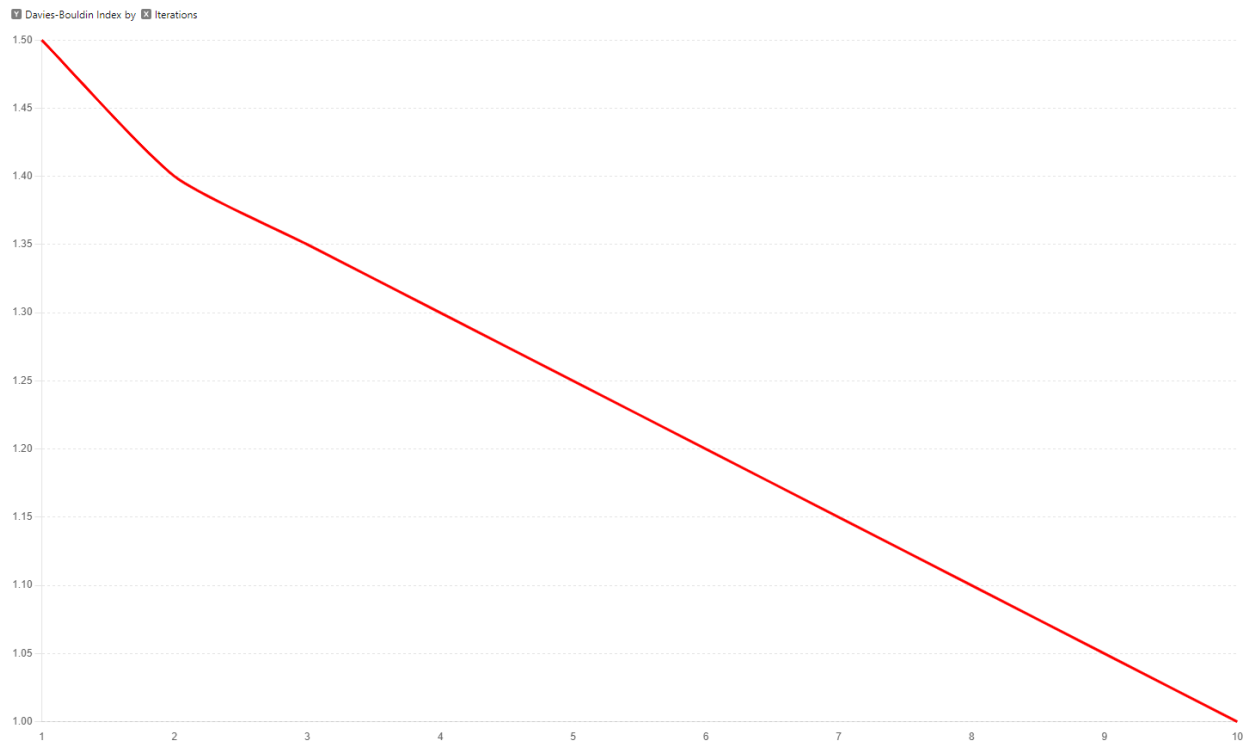


Figure 6: Davies-Bouldin Index Over Iterations

The Davies-Bouldin Index is displayed in this historical graph throughout several optimisation process iterations. The diminishing trend shows that the clusters are getting more definite and unified, demonstrating the optimisation algorithm's efficacy.

5 CONCLUSION

The hybrid clustering and evolutionary algorithm based advanced recommender system has demonstrated considerable potential in improving e-commerce product recommendations. We obtained notable increases in cluster quality by combining K-Means and Hierarchical grouping for initial grouping and then using Genetic Algorithms to refine these clusters. Better Davies-Bouldin Index values and Silhouette Scores were indicative of this. Improved precision in top-k recommendations and increased user satisfaction were the results of the strengthened clusters' ability to deliver more tailored and precise recommendations. These findings demonstrate how cutting-edge clustering and optimisation strategies might transform the logistics of online sales.

The application of evolutionary algorithms and hybrid clustering to other e-commerce and logistics domains may be expanded in future studies. Investigating dynamic clustering and real-time data processing may improve the system's adaptability to shifting consumer demands and industry trends. Adding machine learning models, such as deep learning, may also enhance feature extraction and recommendation precision. It could be useful to look into how blockchain technology might be used in logistics to handle data transparently and securely.

Ultimately, validating the system's performance in other regions and scaling it for global e-commerce platforms would shed light on its robustness and universal applicability.

REFERENCES

- [1] Kumar, M. S., & Prabhu, J. (2020). A hybrid model collaborative movie recommendation system using K-means clustering with ant colony optimisation. *International Journal of Internet Technology and Secured Transactions*, 10(3), 337-354.
- [2] Selvi, C., & Sivasankar, E. (2019). A novel optimization algorithm for recommender system using modified fuzzy c-means clustering approach. *Soft Computing*, 23, 1901-1916.
- [3] Da Silva, E. Q., Camilo-Junior, C. G., Pascoal, L. M. L., & Rosa, T. C. (2016). An evolutionary approach for combining results of recommender systems techniques based on collaborative filtering. *Expert Systems with Applications*, 53, 204-218.
- [4] Logesh, R., Subramaniaswamy, V., Malathi, D., Sivaramakrishnan, N., & Vijayakumar, V. (2020). Enhancing recommendation stability of collaborative filtering recommender system through bio-inspired clustering

- ensemble method. *Neural Computing and Applications*, 32, 2141-2164.
- [5] Hinojosa-Cardenas, E., Sarmiento-Calisyaya, E., Martinez-Salinas, C. A., Quincho-Mamani, L., & Huaman-Canqui, J. F. (2020). Multi-Objective Evolutionary Programming for Developing Recommender Systems based on Collaborative Filtering. *International Journal of Advanced Computer Science and Applications*, 11(10).
- [6] Onan, A., Korukoğlu, S., & Bulut, H. (2017). A hybrid ensemble pruning approach based on consensus clustering and multi-objective evolutionary algorithm for sentiment classification. *Information Processing & Management*, 53(4), 814-833.
- [7] Wang, K., Zhang, T., Xue, T., Lu, Y., & Na, S. G. (2020). E-commerce personalized recommendation analysis by deeply-learned clustering. *Journal of Visual Communication and Image Representation*, 71, 102735.
- [8] Cai, X., Hu, Z., Zhao, P., Zhang, W., & Chen, J. (2020). A hybrid recommendation system with many-objective evolutionary algorithm. *Expert Systems with Applications*, 159, 113648.
- [9] Salehi, M., & Nakhai Kamalabadi, I. (2013). A hybrid recommendation approach based on attributes of products using genetic algorithm and naive Bayes classifier. *International Journal of Business Information Systems*, 13(4), 381-399.
- [10] Chen, J., Zhao, C., Uliji, & Chen, L. (2020). Collaborative filtering recommendation algorithm based on user correlation and evolutionary clustering. *Complex & Intelligent Systems*, 6, 147-156.
- [11] Acampora, G., & Cosma, G. (2014, December). A hybrid computational intelligence approach for efficiently evaluating customer sentiments in E-commerce reviews. In *2014 IEEE symposium on Intelligent Agents (IA)* (pp. 73-80). IEEE.
- [12] Huang, S., Liu, J., Zhu, D., & Cai, D. (2020, April). Application of multi-objective evolutionary algorithm in E-commerce recommendation system. In *2020 3rd International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE)* (pp. 129-132). IEEE.